AN OVERVIEW OF THE LONGITUDINAL ADMINISTRATIVE DATABANK: EVALUATION OF REPRESENTATIVITY

Daniela Ravindra and Linda Standish
Statistics Canada

ABSTRACT

In 1994 Statistics Canada created an 11-year (1982-1992) Longitudinal Administrative Databank (LAD). The LAD was obtained by taking a 1% Bernoulli sample of people with Social Insurance Numbers present on the tax family file (T1FF). The T1FF, and therefore the LAD, are based on information obtained from the annual personal income tax forms filed with Revenue Canada Taxation.

This paper presents some details of the LAD sample, focusing on the representativity of the sample. To assess the representativity of the LAD sample, the sample-based estimates are compared with those of the T1FF population. Estimates were calculated with accompanying coefficients of variance to see if the resulting range includes the actual values. Comparisons by family composition and by age distribution also showed that the LAD sample is a reliable representation of the entire population. Different income variables were tested for representativity to assess large sample sizes (tight confidence limits) and small sample sizes (large confidence limits).

The LAD variables were treated with a process known as capping, whereby outlying observations were changed to the value of a certain threshold. In most cases this threshold was taken at the 99.5 percentile for the variable in question, but sometimes other considerations were taken into account. For incomes which can be large and negative, the lower 0.5 percentile was used for capping. Capping introduces a bias into estimates. The extent of this bias depends on what is being estimated. Therefore the effects of capping had to be taken into consideration when tests of representativity were performed on capped variables.

Although the strength of the representativity for each variable tested was distinct, the test results indicate that the 11 annual samples are acceptable. Inevitably, the sample size and the nature of the distribution of the variable in question will affect the degree of representativity. In summary, the Longitudinal Administrative Databank sample is a reliable and accurate representation of the population.

1. Introduction

The Tax Family File (T1FF) System was introduced in 1982. It is based entirely on information contained in the annual income tax return (T1 form) obtained directly from Revenue Canada Taxation. The T1 file is individual-based. It contains personal information as well as information on dependents. This information is used as the basis for the matching and imputation procedures that produce the T1FF. The T1FF contains both individual and family information. Since the T1

file obtained from Revenue Canada taxation is an administrative file designed to provide the information needed to administer the tax system, it is susceptible to changes in tax policy. As such, the system is reviewed each year in order to implement changes dictated by the tax policy of the year.

The LAD sample was obtained by taking a 1% Bernoulli sample of tax filers and their dependents with a SIN from the T1FF. Taxfilers and dependents with Social Insurance Numbers (SINs) were selected on the basis of random numbers which used the SIN as a seed. A given individual will have the same random number for every year. All individuals which were selected in the first year will be selected in all subsequent years based on these random numbers. The larger a family, the higher the chance that one or more of its members will be selected. Therefore family information, such as family composition, number of people in the family with SIN and family geography may be repeated if more than one member of the same family is selected in the sample.

To each person or family included in the sample a weight is attached. The weight is the inverse of the inclusion probability. For individuals with SINs, the weight is simply 100. For families the weight is more complicated: it is the inverse of the number of people with SINs in the family.

The T1FF contains two kinds of records: filers (all have SINs) and imputed records which are made up of dependant children and spouses (some have SINs). All tax filers with a SIN are equally likely to be selected into the sample. Imputed children and some of the imputed spouses are not eligible for selection since there are no SINs on their records.

The income variables contained in the LAD file can be classified into three categories: individual variables; family variables; and non-child variables. Family variables correspond to the family of the selected individuals (e.g. family composition, family size), while individual variables correspond to the individuals selected into the sample. Non-kid variables correspond to the family members who are not regarded as children.

Annual cross-sectional estimates from the 1% LAD sample can be compared directly with the 100% T1FF population. This is a rare statistical opportunity, having both the population and the sample available.

The overview and analysis that follow are based on a 1% LAD sample.

1.1 Overview of the LAD Sample

During the 1982-1992 period, the total T1FF population of Canada increased steadily. If the LAD sample were to be representative of the total population, it had to increase as well. In fact the LAD sample went from 161,460 units in 1982 to 202,390 units in 1992. This represented an increase of 25.3% which compares very well to the 25.5% increase in the T1FF population with SINs during the same period.

Year	Sample Size	Sample as % of Population
1982	161,460	1.0001
1983	162,750	1.0016
1984	164,350	0.9955
1985	166,270	1.0016
1986	175,710	1.0011
1987	178,310	1.0002
1988	184,000	1.0006
1989	190,660	0.9999
1990	196,050	1.0006
1991	198,740	0.9995
1992	202,390	0.9992

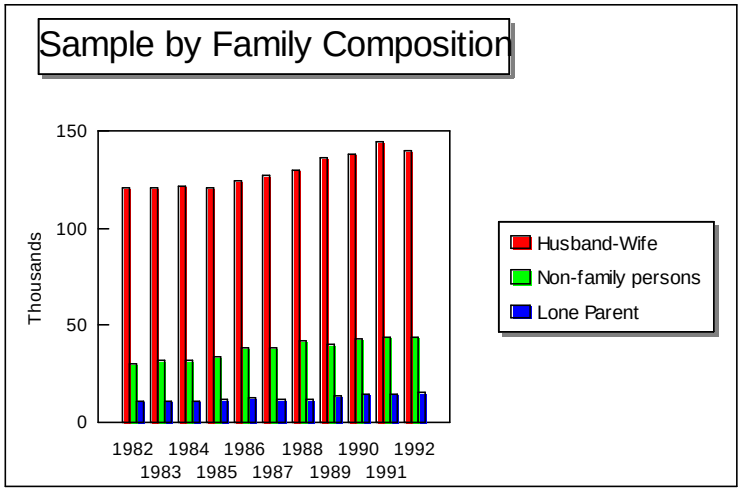
The table above shows that the LAD sample for each year of the 11-year period was extremely close or equal to the 1% mark.

A comparison between the total population represented in the LAD sample and that of the Census of population showed that, for the entire 11-year period, LAD covered over 90% of the Census population estimates. The coverage fluctuated between 91% and 95%. It stayed at 95% for the last three years of the period.

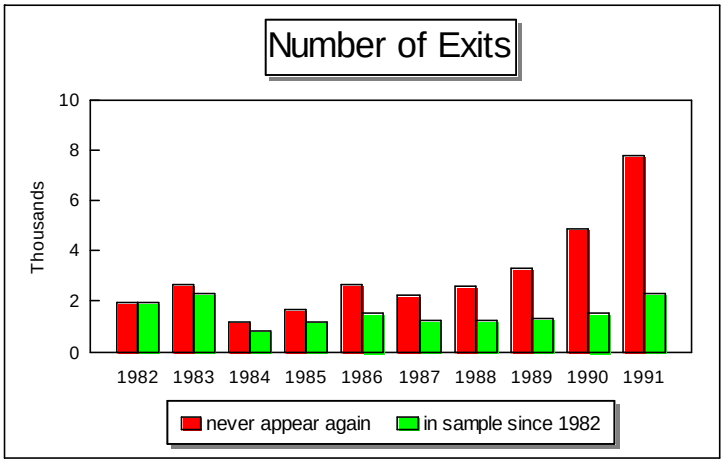
Tax Year	Population Represented in the LAD Sample ('000)	Population Estimates tax year +1 ('000)	LAD Population/ Population Estimate
1982	23,619	24,787	0.95
1983	23,472	24,978	0.94
1984	23,728	25,165	0.94
1985	23,833	26,204*	0.91
1986	24,517	26,550	0.92
1987	24,833	26,895	0.92
1988	25,155	27,379	0.92
1989	25,909	27,791	0.93
1990	26,577	28,120*	0.95
1991	27,007	28,542	0.95
1992	27,534	28,941	0.95

* denotes Census population estimates

The units in the sample belong to one of the three family types: husband-wife, single parent and non-family persons. By far, the majority of the LAD units were people belonging to husband-wife families. The proportion of husband-wife and that of lone-parent families increased steadily, while the proportion of non-family persons experienced slight up and down fluctuations.

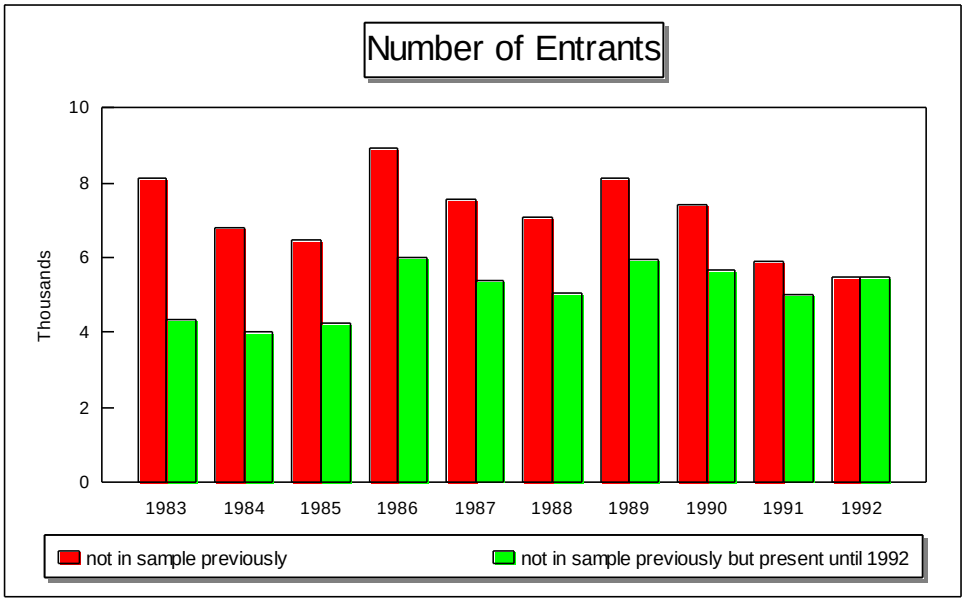


Every year, units exit the sample and new units enter to replace them. There are a number of reasons for units leaving the sample: people die, people stop filing income tax returns or do not file for a period of time only to return to filing later, and people move out of Canada. It is important to note that if people who did not file for a number of years start filing again they will be selected in the sample as soon as they become filers.



The number of exits that never returned to the sample again increased slightly over the years. However, as a percentage of the annual sample it remained small. There is still an opportunity for these individuals to re- appear in the sample subsequently.

The number of entrant units that were not in the sample previously increases twice, in 1986 and 1989.



The 1986 tax year was the first year when filers were eligible for the Federal Sales Tax Credit. As a result there was an increase in the number of people filing. Then, in 1989 more people started filing tax returns in order to apply for the Goods and Services Tax credit that was being introduced the following year. Both these credits are available to people exclusively through the tax system, and do not depend on tax liability.

2. Capping

The LAD income and deduction variables capped, whereby outliers in a particular distribution were changed to the value of a certain threshold. Both positive and negative outliers were capped. The threshold was chosen to be the 99.5 percentile for the variable in question for positive values. For negative outliers, the caps were the 0.5 percentile.

There were two main reasons for capping the LAD variables:

- some income variables are highly skewed with a large number of outliers; capping dealt with the possibility of over-estimation due to an outlier appearing in the LAD sample.

- capping, in part, addresses confidentiality concerns since the process reduces the likelihood of dominance by an outlier in an aggregate.

Capping has an impact on the estimates of aggregates obtained from the capped sample. Within a single variable only the half of one percent at one tail end (or both if the variable's values can be negative) will be altered. However, there are two situations in which capping may have a compounding effect on a given variable:

- when the variable is a family income figure; in this case data for more than one individual in the family may be capped; the likelihood of this situation is small.

- when a variable is a sum of other component variables; in these cases capping was implemented on the individual components before the summation was performed.

In both these cases there is greater distortion than in the case of a single variable.

3. Representativity

Whenever a sample is taken from a population there will always be some discrepancy between the distribution of the sample and that of the population. The representativity of the sample measures how well the sample reflects the population and gives an indication of the accuracy and reliability of the estimates.

The following section will present the results of testing for a number of variables. These variables have been chosen to represent both large sample sizes (tight confidence intervals) and small sample sizes (large confidence intervals). All the data used in the representativity tests were obtained from

the 1992 LAD sample.

3.1 Estimated Coefficients of Variation

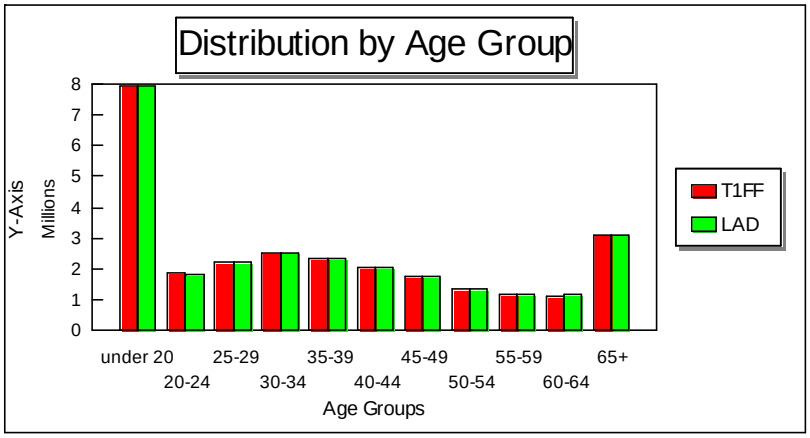
A measure of the reliability of the estimates calculated from sample data is the estimated coefficient of variation which will be denoted by cv. The cv measures the variability of the estimate. The lower the cv, the more reliable the estimate. At Statistics Canada estimates with a cv of 16.5% or less are considered reliable, while estimates with a cv greater than 33.3% are not considered reliable. Estimates with coefficients of variation between 16.5% and 33.3% are treated with caution.

3.2 Confidence Intervals

Using the estimate and the cv one can compute a confidence interval at a specified confidence level. A 95% confidence interval was used for the purpose of calculating confidence intervals for examining the LAD estimates.

3.3 Population by Age Group

The graph below depicts the comparison between the distribution by age group of the T1FF population and that of the LAD sample. The differences are almost negligible. All population counts fall within the confidence intervals calculated from the sample. Table 1 and Table 2 in the Appendix illustrate the distribution for the male and female categories as well as the total. It can be seen that for these categories the estimates are as reliable as the ones for the total. The coefficients of variation are well below the accepted limit of reliability. As expected, the coefficients of variation are larger for the age groups with the smaller sample sizes (i.e. age groups <20 and 60-64).



3.4 Distribution by Age Group and Family Type

Husband-wife families represent the largest proportion (70.4% in 1992) of all family types. Some of them have both parents and children, some are made up of two married or common-law people. Lone parent families must have at least one child. Separate estimates were obtained for the number of children and for the number of parents for these two family types. All these estimates fall within the 95% confidence interval calculated from the sample. The size of the confidence intervals differ substantially, mostly due to the differences in the coefficients of variation. A few of the coefficients of variation are higher than the accepted threshold of reliability. However, in all of these cases the sample sizes are small: 2,500 for children between 50 and 54, 1,500 for children between 55 and 64 and 200 for children over 60 to 64 years old. Only the last category has a coefficient of variation that indicates the data are unreliable.

The results are similar for the lone parent family data. Again, for a few age categories for which the sample sizes are small the coefficients of variation indicate that the data should be treated with caution. For most, however, the coefficients of variation are well below the accepted limit of reliability.

3.5 Population by Marital Status

A comparison of the population by marital status reveals that the differences between the sample and the population are almost negligible. Not only are the values for the totals very similar, but even when the population and the sample are broken down by gender the two values are compatible. The marital status variable is an example of a variable for which the sample size is large, so only small variations between the population and the sample are expected.

The 100% estimates fall within the 95% confidence intervals generated from the sample. The sizes of the confidence intervals differ due to the differences in the coefficients of variation. As expected, the variance was smallest for the variables with the larger populations, i.e. the married and the single groups. All coefficients of variation were within the reliable limit (i.e. <16.5%).

3.6 Total Income Distribution by Age Group for Males

Table 7 in the Appendix gives a distribution of total income by age group for males. The magnitudes of the coefficients of variation indicate that all points of data are reliable. However, there are two instances when the population aggregate falls slightly outside the confidence interval generated from the sample. In both cases the T1FF value is smaller than the sample estimate.There could be two reasons for this: records with large income values are over-represented in the LAD sample; or components of the total income have been capped and then aggregated so the effect of capping was compounded.

3.7 Components of Employment Income

Employment income is the sum of wages, salaries and commissions and self employment income. Over 93% of the employment income is made up of wages, salaries and commissions. The table below contains a comparison of estimates, coefficients of variation and confidence intervals for the two components of employment income. In both cases the T1FF aggregate falls within the 95% confidence interval generated from the sample.

The Bernoulli sampling scheme does not give a fixed sample size. This leads to a larger variance for estimates of aggregates and means, as opposed to the simple random sampling which gives a fixed sample size. Also there is a difference in the size of the population from variable to variable since only the non-zero values are taken into consideration. In this case the population for the wages and salaries variable is much larger than that for the self-employment income. Therefore the estimates derived for self-employment incomes are less reliable than those for wages and salaries.

Variable	**T1FF Aggregate ('000,000)**	**LAD Estimate ('000,000)**	**Confidence Interval (1,000,000)**	**CV %**
Wages and salaries	257,447	251,450	249,730-253,180	0.35
Self-employment	18,147	17,801	17,232-18,371	1.63

3.8 Comparison of Medians

Capping should not affect median values of single variables as it deals with outlier values. In order to test this, medians of employment income for the three family types were compared. As the table shows below, the medians obtained from the LAD were identical to the ones from the T1FF.

Family Type	**T1FF median of Employment Income**	**LAD median of Employment Income**
Husband-Wife	47,500	47,500
Lone-Parent	21,700	21,700
Non-family Person	15,400	15,400

4. Conclusions

A number of tests were performed on variables to determine if the LAD sample represented well the T1FF population. Some of the results were:

- over the 11 years the sampling was the expected 1% of the T1FF population of records with SINs;

- individual as well as family demographic variables are well represented in the sample, a good reflection of the population;

- estimates of income components are representative of distributions of the population, but the population aggregates are at times outside the 95% confidence intervals calculated from the sample;

- capping had an impact on the combined income variables; and

- medians calculated from the LAD were identical to the T1FF ones for single component variables.

The above highlights show that, although the LAD provides an accurate representation of the T1FF, there are a few areas of disagreement. The LAD estimates of rare populations, such as widowed individuals, or individuals in the <20 age group were not representative of the T1FF. Also some aggregate amounts are distorted by capping.

When the 10% LAD is created there will be no capping done. Capping of all income variables will be removed and the 10% sample recreated for the 11 years. This way income variables can be studied longitudinally without having to take into account the alterations that capping may have introduced. If aggregates from LAD are unreliable, they can be obtained directly from the 100% file (T1FF) for cross-sectional studies, rather than aggregating the values in the LAD. Thus the LAD will concentrate on longitudinality, not on cross-sectional aggregates.

Data presented in this paper are for the national level. However, the LAD sample can support comparisons at the provincial level as well. Some preliminary studies have been done (Saleh and Demnati) and they show that even though data for larger provinces well represent the T1FF, even the small provinces have large enough samples to support research for most variables.

5. Future Considerations

LAD provides a rich source of data for researchers and policy makers.Presently, it covers a period of 11 years and is the longest longitudinal file at Statistics Canada. The table below presents a summary of the minimum population size required for a given CV and a given sampling proportion:

MINIMUM POPULATION SIZE

Sampling proportion	cv = 10%	cv = 20%	cv = 30%
1%	9,900	2,475	1,100
5%	1,900	475	211
10%	900	225	100

The minimum population size is fairly small at the 10% sampling level, even when a cv of 10% is sought. Therefore estimates of counts from rare populations (e.g. fishing income), or estimates at small area levels will be obtainable once the 10% sample is available.

Future considerations may include:

- Because capping will be eliminated, the representativity tests, especially for aggregates of incomes which are subjected to outliers with large magnitude should be performed again.

- Maintaining the sample at a 10% level. Until now the sample was maintained at 1% due to storage considerations on the available hardware but it is in the process of being expanded to 10%.

References

FEDER, M. and DEMNATI, A. (1994) User's Guide to LAD Cross-Sectional Estimation, Statistics Canada, unpublished

HARRIS, S. and LUCACIU, D. (1994) An Overview of the T1FF Creation, Statistics Canada, unpublished

LUCACIU, D. and ST-LOUIS, A. (1994) An Overview of the LAD Sample, Statistics Canada, unpublished

SALEH, M. and DEMNATI, A. (1994) Representativity of the LAD Sample, Statistics Canada, unpublished

SALEH, M. and DEMNATI, A. (1994) The Effects of Capping, Statistics Canada, unpublished

Appendix

Table 1: Distribution by Age Group for Males

Age Groups	Males			
	T1FF	LAD	Confidence Interval	CV %
Under 20	377,900*	392,640	380,230-405,040	1.58
20-24	907350	920,270	901,220-939,310	1.03
25-29	1,085,530	1,072,730	1,052,200-1,093,260	0.96
30-34	1,223,730	1,230,330	1,208,370-1,252,280	0.89
35-39	1,156,440	1,166,900	1,145,500-1,188,300	0.92
40-44	1,026,290	1,025,500	1,005,420-1,045,570	0.98
45-49	878,630	885,300	866,640-903,950	1.05
50-54	682,560	672,000	635,740-688,260	1.21
55-59	588,670	584,000	568,870-599,130	1.30
60-64	566,590	577,400	562,320-592,480	1.31
65+	1,331,270	1,340,600	1,317,650-1,363,540	0.86
Total	9,824,960	9,867,670	9,805,400-9,929,930	0.32

* outside the confidence interval

Note: Gender is not available for imputed children. Thus, the number of males < 20 is incomplete.

Table 2: Distribution by Age Group for Females

Age Group	Females			
	T1FF	LAD	Confidence Interval	CV %
Under 20	369,610	372,570	360,510-384,620	1.62
20-24	935,950	924,510	905,460-943,550	1.03
25-29	1,137,040	1,131,700	1,110,680-1,152,720	0.93
30-34	1,272,320	1,264,800	1,242,560-1,287,040	0.88
35-39	1,200,840	1,195,500	1,173,900-1,217,100	0.90
40-44	1,058,790	1,059,600	1,039,260-1,079,940	0.96
45-49	889,790	886,500	867,910-905,090	1.05
50-54	691,690	687,000	670,640-703,360	1.19
55-59	598,840	592,600	577,420-607,780	1.28
60-64	577,470	577,800	562,810-592,790	1.30
65+	1,785,580	1,781,600	1,755,210-1,807,990	0.74
Total	10,517,580	10,474,170	10,410,220-10,538,120	0.31

Note: Gender is not available for imputed children. Thus, the number of females <20 is incomplete.

Table 3: Distribution by Age Group for Total Population

Age Group	Total			
	T1FF	LAD	Confidence Interval	CV %
Under 20	7,923,130	7,927,390	7,862,720-7,992,060	0.41
20-24	1,854,930	1,852,380	1,825,370-1,879,380	0.73
25-29	2,224,550	2,206,240	2,176,850-2,235,630	0.67
30-34	2,496,410	2,495,280	2,464,020-2,526,530	0.63
35-39	2,358,740	2,364,390	2,333,250-2,395,530	0.64
40-44	2,084,770	2,085,100	2,056,520-2,113,680	0.69
45-49	1,768,690	1,771,900	1,745,420-1,798,380	0.74
50-54	1,374,260	1,359,100	1,335,840-1,382,360	0.85
55-59	1,187,540	1,176,600	1,155,170-1,198,030	0.91
60-64	1,144,060	1,155,200	1,133,930-1,176,470	0.92
65+	3,116,840	3,122,200	3,087,230-3,157,170	0.56
Total	27,533,910	27,515,770	27,406,900-27,624,650	0.20

Table 4: Distribution by Marital Status for Males

Marital Status	Males			
	T1FF	LAD	Confidence Interval	CV %
Single	2,882,210	2,904,370	2,870,490-2,938,240	0.58
Married	5,680,090	5,707,700	5,660,370-5,755,030	0.41
Common Law	274,770	274,200	264,030-284,370	1.85
Separate/ Divorced	760,020	745,500	728,380-762,620	1.15
Widowed	227,870	234,700	225,100-244,300	2.05
Total	9,824,960	9,866,470	9,804,200-9,928,730	0.32

Table 5: Distribution by Marital Status for Females

Marital Status	Females			
	T1FF	LAD	Confidence Interval	CV %
Single	2,465,990	2,453,670	2,422,540-2,484,800	0.63
Married	5,669,090	5,646,500	5,599,7305,693,270	0.41
Common Law	276,910	282,200	271,890-292,510	1.83
Separate/ Divorced	1,042,510	1,082,800	1,052,970-1,012,620	0.98
Widowed	1,063,080	1,057,800	1,037,340-1,078,260	0.97
Total	10,517,580	10,472,970	10,409,030-10,536,920	0.31

Table 6: Distribution by Marital Status for Total Population

Marital Status	Total			
	T1FF	LAD	Confidence Interval	CV %
Single	12,539,570	12,531,970	12,454,420-12,609,530	0.31
Married	11,349,180	11,356,400	11,289,850-11,422,950	0.29
Common Law	551,680	556,600	542,120-571,080	1.30
Separate/ Divorced	1,802,530	1,778,300	1,751,840-1,804,760	0.74
Widowed	1,290,950	1,292,500	1,269,900-1,315,100	0.87
Total	27,533,910	27,515,770	27,406,850-27,624,690	0.20

Table 7 Total Income Distribution of Males by Age Group

Thousands of dollars	Males <25 years of age			
	T1FF	LAD	Confidence Interval	CV %
under 5.0	349,300	353,300	341,480-365,120	1.67
5.0 - 9.9	361,240	363,300	351,310-375,290	1.65
10.0-14.9	210,930	212,600	203,430-221,770	2.16
15.0-19.9	135,750	140,300	132,850-147,750	2.66
20.0-24.9	91,530	94,400	88,290-100,510	3.24
25.0-34.9	88,960	92,000	85,960-98,040	3.28
35.0-49.9	32,860	34,000	30,330-37,670	5.40
50.0-74.9	4,830	5,500	4,040-6,360	13.30
75.0-99.9	350	400	2-800	49.70
100.0+	220	100	-100-300	99.50
Total	1,275,970	1,295,900	1,273,250-1,318,550	0.87

Thousands of dollars	Males 25-34 years of age			
	T1FF	LAD	Confidence Interval	CV %
under 5.0	141,190	142,400	134,910-149,980	2.63
5.0 - 9.9	230,250	233,500	223,890-243,110	2.06
10.0-14.9	219,570	217,700	208,420-226,980	2.13
15.0-19.9	237,780	233,100	223,500-242,700	2.06
20.0-24.9	251,910	249,900	239,970-258,830	1.99
25.0-34.9	478,360	476,000	462,280-489,720	1.44
35.0-49.9	473,260	467,000	453,400-480,600	1.46
50.0-74.9	197,190	195,600	186,800-204,400	2.25
75.0-99.9	22,430*	26,000	22,790-29,210	6.17
100.0+	12,780	12,400	10,190-14,600	8.90
Total	2,264,710	2,253,600	2,223,750-2,283,450	0.66

* outside confidence interval

Thousands of dollars	Males 35-44 years of age			
	T1FF	LAD	Confidence Interval	CV %
under 5.0	104,990	105,500	99,070-111,930	3.05
5.0 - 9.9	159,170	159,000	151,0780-166,920	2.49
10.0-14.9	140,330	137,300	129,930-144,670	2.68
15.0-19.9	148,090	149,200	141,530-156,870	2.57
20.0-24.9	162,980	161,400	153,420-169,380	2.47
25.0-34.9	362,630	356,000	344,130-367,870	1.67
35.0-49.9	514,890	518,800	504,470-533,130	1.38
50.0-74.9	404,000*	420,000	407,108-432,890	1.53
75.0-99.9	77,770	82,800	77,070-88,530	3.46
100.0+	54,870	51,200	46,790-55,810	4.39
Total	2,129,720	2,141,300	2,112,210-2,170,400	0.68

* outside confidence interval

Thousands of dollars	Males 45-54 years of age			
	T1FF	LAD	Confidence Interval	CV %
under 5.0	77,640	82,900	77,180-88,620	3.45
5.0 - 9.9	118,220	112,800	106,130-119,470	2.96
10.0-14.9	97,030	95,200	89,070-101,330	3.22
15.0-19.9	95,910	93,200	87,140-99,270	3.26
20.0-24.9	101,480	100,200	93,900-106,500	3.14
25.0-34.9	227,200	225,400	215,950-234,850	2.10
35.0-49.9	335,720	343,400	331,740-355,060	1.70
50.0-74.9	320,680	321,200	309,920-332,480	1.76
75.0-99.9	81,310	84,600	78,810-90,390	3.42
100.0+	63,240	60,300	55,410-65,190	4.05
Total	1,518,440	1,519,200	1,494,680-1,543,720	0.81

Thousands of dollars	Males 55-64 years of age			
	T1FF	LAD	Confidence Interval	CV %
under 5.0	61,150	61,000	56,120-65,880	4.00
5.0 - 9.9	119,100	121,300	114,370-123,230	2.86
10.0-14.9	102,170	96,800	90,620-102,980	3.19
15.0-19.9	97,960	98,000	90,770-104,230	3.18
20.0-24.9	102,270	99,700	93,420-105,980	3.15
25.0-34.9	205,900	214,300	205,090-223,510	2.15
35.0-49.9	216,390	221,000	211,640-230,360	2.12
50.0-74.9	142,540	144,000	136,450-151,550	2.62
75.0-99.9	41,510	41,200	37,160-54,240	4.90
100.0+	40,430	37,200	33,370-41,030	5.15
Total	1,129,430	1,134,500	1,113,320-1,155,680	0.93

Thousands of dollars	Males 65+ years of age			
	T1FF	LAD	Confidence Interval	CV %
under 5.0	28,410	29,000	26,020-31,980	5.14
5.0 - 9.9	76,700	76,700	71,220-82,180	3.57
10.0-14.9	356,770	359,100	347,180-371,020	1.66
15.0-19.9	238,510	237,600	227,900-247,300	2.04
20.0-24.9	174,120	177,000	168,630-185,370	2.36
25.0-34.9	217,140	216,800	207,530-226,070	2.14
35.0-49.9	130,150	131,500	124,280-138,720	2.74
50.0-74.9	63,700	68,800	63,580-74,020	3.79
75.0-99.9	17,440	16,700	14,130-19,270	7.70
100.0+	20,430	18,100	15,420-20,780	7.40
Total	1,323,370	1,331,300	1,308,410-1,354,190	0.86

Table 8: Total Income Distribution of Females by Age Group

Thousands of Dollars	Females < 25 years of age			
	T1FF	LAD	Confidence Interval	CV %
under 5.0	393,220	387,600	375,220-399,980	1.60
5.0 - 9.9	388,900	392,000	379,550-404,450	1.59
10.0-14.9	229,700	227,500	218,010-236,990	2.09
15.0-19.9	130,070*	122,400	115,440-129,360	2.84
20.0-24.9	73,530	76,000	70,510-81,490	3.61
25.0-34.9	55,910	55,200	50,520-59,880	4.23
35.0-49.9	11,620*	9,500	7,560-11,440	10.20
50.0-74.9	1,100	1,200	510-1,890	18.70
75.0-99.9	140	100	-100-300	99.50
100.0+	90*	X	X	X
Total	1,284,270	1,271,500	1,249,070-1,293,930	0.88

Thousands of Dollars	Females 25-34 Years of age			
	T1FF	LAD	Confidence Interval	CV %
under 5.0	389,190	393,900	381,440-406,360	1.58
5.0 - 9.9	300,700	301,700	290,790-312,610	1.81
10.0-14.9	353,350	349,000	337,250-360,750	1.68
15.0-19.9	320,430	314,800	303,640-325,960	1.77
20.0-24.9	274,260	278,400	267,800-288,900	1.89
25.0-34.9	398,480	395,200	382,690-407,710	1.58
35.0-49.9	232,130	233,000	223,400-242,610	2.06
50.0-74.9	53,570	51,800	47,280-56,320	4.37
75.0-99.9	5,300*	3,900	2,660-5,140	15.90
100.0+	2,790	2,300	1,350-3,250	20.70
Total	2,330,190	2,324,000	2,293,690-2,354,110	0.65

* outside confidence interval

Thousands of Dollars	Females 35-44 years of age			
	T1FF	LAD	Confidence Interval	CV %
under 5.0	338,540	340,300	328,730-351,870	1.70
5.0 - 9.9	244,760	242,900	233,100-252,690	2.02
10.0-14.9	275,150	272,200	261,820-282,580	1.90
15.0-19.9	247,690*	235,600	225,950-245,260	2.05
20.0-24.9	222,620	227,700	218,210-237,190	2.08
25.0-34.9	381,910	378,500	366,260-390,740	1.62
35.0-49.9	280,330	289,200	278,500-299,900	1.85
50.0-74.9	130,030	130,100	122,920-137,280	2.76
75.0-99.9	15,160	13,500	11,200-15,800	8.53
100.0+	9,010	7,900	6,130-9,670	11.20
Total	2,145,190	2,137,900	2,108,830-2,166,970	0.68

Thousands of Dollars	Females 45-54 years of age			
	T1FF	LAD	Confidence Interval	CV %
under 5.0	216,050	212,300	203,140-221,410	2.15
5.0 - 9.9	203,800	209,800	200,700-218,910	2.17
10.0-14.9	184,250	181,800	173,320-190,280	2.33
15.0-19.9	152,730	150,800	143,080-158,520	2.56
20.0-24.9	137,750	136,600	129,240-143,960	2.69
25.0-34.9	237,640	236,300	226,630-245,970	2.05
35.0-49.9	178,340	179,800	171,360-188,240	2.35
50.0-74.9	99,580	97,900	91,670-104,130	3.18
75.0-99.9	13,080	11,100	9,000-13,200	9.44
100.0+	7,400*	5,700	4,200-7,200	13.20
Total	1,430,620	1,422,100	1,398,400-1,445,800	0.83

* outside confidence interval

Thousands of Dollars	Females 55-64 years of age			
	T1FF	LAD	Confidence Interval	CV %
under 5.0	232,840	234,700	225,220-244,180	2.02
5.0 - 9.9	229,060	232,900	223,320-242,480	2.06
10.0-14.9	156,080	154000	146,200-161,800	2.53
15.0-19.9	104,670	102,800	96,420-109,180	3.10
20.0-24.9	83,050	85,000	79,200-90,800	3.41
25.0-34.9	125,470	124,700	117,670-131,730	2.82
35.0-49.9	77,210	76,700	71,190-82,210	3.59
50.0-74.9	37,600	37,200	33,360-41,040	5.16
75.0-99.9	7,390	7,300	5,600-9,000	11.60
100.0+	5,600	5,000	3,590-6,410	14.10
Total	1,058,950	1,060,300	1,039,840-1,080,700	0.96

Thousands of Dollars	Females 65+ years of age			
	T1FF	LAD	Confidence Interval	CV %
under 5.0	157,930	158,900	151,340-166,460	2.38
5.0 - 9.9	303,680	304,900	249,030-315,770	1.78
10.0-14.9	685,010	689,400	672,880-705,920	1.20
15.0-19.9	223,980	221,500	212,140-230,860	2.11
20.0-24.9	127,640	122,700	115,730-129,670	2.84
25.0-34.9	138,010	141,000	133,530-148,470	2.65
35.0-49.9	74,290	73,200	67,820-78,580	3.68
50.0-74.9	32,280	34,200	30,520-37,880	5.38
75.0-99.9	8,090	6,900	5,250-8,550	12.00
100.0+	7,600*	4,300	2,990-5,610	15.20
Total	1,758,530	1,757,000	1,730,780-1,783,220	0.75

* outside confidence interval